

Telephone vs Face-to-Face Interviewing in a Community Psychiatric Survey

ABSTRACT

This study compared telephone with face-to-face interviewing in a community psychiatric survey. Two groups of women were investigated, Holocaust survivors and Europe-born respondents who were in prestate Israel during World War II. Both were administered the Psychiatric Research Interview Demoralization Scale and a short item scale investigating World War II experiences. Results showed a high compliance rate to the telephone mode. The subjects' scores in the two modes were highly correlated. Telephone interviewing seems to be a reliable and efficient method in areas with a well-developed network of subscribers. (*Am J Public Health.* 1993;83:896-898)

Shmuel Fenig, MD, Itzhak Levav, MD, Robert Kohn, MD, and Nava Yelin, BA

Introduction

Psychiatric epidemiology surveys usually rely on face-to-face interviews, while telephone questionnaires are seldom used. Yet, telephone interview-based surveys have numerous advantages. Their cost is one fourth to one half of that of face-to-face interviews.¹⁻³ Their use permits selecting populations based on broader sampling frameworks, and they may represent the most feasible method to conduct surveys that include relatives of probands residing in geographically distant places.⁴ The telephone also increases completeness by enabling access to respondents who are hard to reach because of shifting work schedules, residence in dangerous locales, or resistance to face-to-face interviews. Finally, the partial anonymity granted by the telephone may increase the validity of responses by reducing the embarrassment involved in responding to emotionally or socially loaded questions in face-to-face situations.

The telephone, however, is not free from shortcomings. Bias may arise from the exclusion of potential respondents without telephone service or with unlisted numbers whenever telephone directories are used for sample selection. Some studies suggest that telephone-based interviews lead to missing information regarding sensitive items^{2,5-7} and to more positive responses regarding psychiatric symptoms.⁸ Others, in contrast, find no differences between the two modes regarding compliance, reliability of replies to sensitive questions, or outcome measures.⁹⁻¹²

Most previous studies, however, have not based their conclusions on repeated interviews but on random subsamples of respondents who were administered one of the two methods. Also, these reports usually examine the comparability of responses to single items and not to larger question sets. A third weakness of prior reports using psychiatric diagnostic interviews is the prolonged interval between modes, ranging from 3 to 19 months.^{4,13}

Unaddressed methodological issues still remain, although the literature published thus far leans favorably toward the use of the telephone as a viable interview modality. Can favorable findings still be demonstrated when the time elapsed between interview waves is brief, when complex question sets are administered, and when especially sensitive populations are tapped?

Methods and Procedures

The sample was composed of two groups. The index group included women who experienced the Holocaust in World War II. The comparison group was made up of women who resided in prestate Israel during this period and who possessed the same sociodemographic characteristics as those in the index group.

The prospective respondents were selected from the computerized records of the General Federation of Labor, with an affiliation of over 80% of the residents of Tel Aviv. One hundred seventy women met the sociodemographic criteria: being married, born in Poland between 1919 and 1924, and residing in Tel Aviv. Those ($n = 153$) who had telephone service were included in the study.

Demoralization was explored with the aid of the Demoralization Scale of the Psychiatric Epidemiology Research Interview.^{14,15} This scale had been tested in Israel for internal consistency reliability and validity in heterogeneous community and psychiatric samples.¹⁶ Items were scored on a four-point Likert scale. The

Shmuel Fenig is with the Shalvata Mental Health Center, Hod Hasharon and Sackler School of Medicine, University of Tel Aviv, Tel Aviv, Israel. Itzhak Levav is with WHO/Pan American Health Organization, Washington, DC. Robert Kohn is with Brown University, Providence, RI. Nava Yelin is with the Kupat Holim Central Offices, Tel Aviv, Israel.

Requests for reprints should be sent to Itzhak Levav, MD, Division of Health Promotion, WHO/Pan American Health Organization, 525 23rd St, NW, Washington, DC 20037.

This paper was accepted December 28, 1992.

Demoralization Scale, a subscale of a larger psychiatric screen, can be used to categorize subjects with or without demoralization.

The degree of victimization was measured by the following items in ascending degree of traumatization: being displaced in the Soviet Union, being confined to a camp in the Soviet Union, staying in a ghetto, being in hiding, being in forced labor, and staying in an extermination camp. The highest degree of victimization was coded, regardless of the length of exposure.

A letter, signed by the chief nurse of the respondent's health center and by the first author, was sent to the individual's home address requesting cooperation with the survey. All respondents were contacted within 2 weeks and were administered the Demoralization Scale, the sociodemographic items, and the items related to respondents' location during World War II.

The face-to-face interviews were conducted by two closely supervised mental health workers. Interrater reliability ($\kappa = .81$ to $.85$; $n = 20$) between the mental health workers and the first author was determined after training. The face-to-face interviewers were blind to the telephone interview results. The time interval between the interviews ranged from 5 to 21 days (10 days on average). A Demoralization Scale value of 1.55 was adopted as a cutoff to differentiate women with and without a degree of demoralization compatible with a psychiatric condition.¹⁶

Results

Eight of the original 153 selected respondents refused to be interviewed over the telephone. Of the 145 telephone respondents, 52.4% were Holocaust survivors, while 47.6% were from the comparison group. The mean age for the sample was 65.6 (SD = 1.9) years. As a result of budgetary constraints, only 115 respondents were randomly selected for the household stage, with an 83% completion rate ($n = 96$). As required by the health center, no attempts were made to convert refusals. No differences were found in age, year of immigration, education, social status, degree of victimization, religiosity, or demoralization between those who did and did not undergo the household stage.

Cronbach's alpha for the Demoralization Scale showed high reliability for both telephone (.94) and face-to-face interviews (.90), with a negligible drop be-

TABLE 1—Subgroup Comparisons of Demoralization Scale Scores between Telephone and Face-to-Face Interviews

	n	Demoralization Scale Score, Mean \pm SD		<i>P</i> ^a	Pearson Correlation
		Telephone	Face-to-Face		
Traumatization					
Holocaust	49	1.84 \pm 0.76	1.45 \pm 0.66	<.001	.79
Non-Holocaust	47	1.15 \pm 0.52	1.06 \pm 0.48	NS	.68
Education, y					
1–8	32	1.80 \pm 0.83	1.47 \pm 0.64	<.003	.72
9–12	49	1.37 \pm 0.68	1.15 \pm 0.56	<.001	.79
12+	15	1.29 \pm 0.57	1.19 \pm 0.66	NS	.87

Note. *P* values are unadjusted. NS = not significant.
^aPaired *t* test.

tween modes. The scores for the two interviewing modes had a high Pearson correlation (.78).

The rates of demoralization were 40% for telephone interviews (mean score = 1.47, SD = 0.77) and 30.2% for face-to-face interviews (mean score = 1.26, SD = 0.61) (*t* test, $P < .001$). Paired *t* tests with Bonferroni correction for multiple comparisons ($P = .001$) showed that only 6 of the 27 items of the Demoralization Scale had significantly different responses between telephone and face-to-face interviews. Without such a correction, 12 items were statistically significant, with the means higher for telephone interviews.

The correlations and differences in means between methods were examined by Holocaust experience and education (see Table 1). Item analyses revealed that 6 of the 27 Demoralization Scale item responses had significantly different means ($P = .001$) between telephone and face-to-face interviews for the Holocaust group; there were no significant differences in the non-Holocaust group. The results (at $P = .05$) comparing the two interview modes showed that 12 and 6 items differed significantly for the respective groups. We also ran a regression model with the change score between telephone and face-to-face interviews as the dependent variable and Holocaust experience, education, and the interaction term as covariates. The difference between the two interviewing modes was accounted for by Holocaust status ($\beta = -.67$, $P = 0.001$) and its interaction with education ($\beta = .05$, $P = .004$); education itself was not significant ($\beta = -.01$).

Discussion

Our results suggest that the telephone can be an excellent screening method; the

scores obtained correlated highly with those elicited in face-to-face interviews. Our results differed from a previous report indicating that telephone respondents tend to provide more "cheerful" responses.⁸ This study also demonstrates a high compliance with telephone interviews. The willingness of Holocaust survivors to discuss aspects of their lives by telephone indicates that this method can be used, with the necessary precautions, even in highly sensitive populations.

The lower mean demoralization scores in all subgroups in the face-to-face interviews, conducted last, can be explained by regression to the mean, by increased sensitivity of the telephone mode, by respondents taking the second evaluation less seriously, by subjects trying to create a more favorable impression on retest,¹⁷ and/or by therapeutic effects among demoralized groups.¹⁸ Unfortunately, the lack of randomization of the order of administration between modes prevents a more conclusive analysis of this problem. Closer scrutiny of the Holocaust group, however, contributes in part to a resolution of this issue. Little difference between the two interview modes was noted for the non-Holocaust group, suggesting an equivalence between methods. In contrast, there was a marked decrease in Demoralization Scale scores for the Holocaust group. One can argue that the initial higher demoralization scores resulted in a greater regression to the mean. Alternatively, the greater psychological distance of the telephone interview may have facilitated symptom expression. A therapeutic effect of the initial interview can be proposed as a third explanation; in view of previous results,¹⁸ it may seem a persuasive explanation.

The brief time lag between interviewing methods is a strength of this study, minimizing change within subjects. Alternatively, it may have increased respondents' memory of earlier responses, an unlikely possibility in view of the differential behavior of the two groups.

Since this study used a psychiatric screening scale, its results cannot necessarily be extrapolated to diagnostic procedures. The literature, however, suggests that the latter can be performed by telephone as well.^{4,13}

Because industrialized nations possess a developed telephone system, this method of interviewing is becoming an attractive and cost-efficient fieldwork resource. Caution should be exerted, however, regarding bias arising from selective exclusion of subscribers or from uneven network development (e.g., inner-city areas or rural zones). □

Acknowledgments

We thank the Kupat Holim clinics that collaborated in the study and the reviewers for their helpful suggestions.

References

1. Weeks MF, Kulka RA, Lessler JT, Whitmore RW. Personal versus telephone surveys for collecting household health data at the local level. *Am J Public Health*. 1983; 73:1389-1394.
2. Groves RM, Kahn RL. *Surveys by Telephone: A National Comparison with Personal Interviews*. New York, NY: Academic Press; 1979.
3. Yaffe R, Shapiro S, Fuchsberg RR, Rohde CA, Corpeno HC. Medical economics survey-methods study: cost effectiveness of alternative survey strategies. *Med Care*. 1978;16:641-659.
4. Paulson AS, Crawe RR, Noyes R, Pfohl B. Reliability of the telephone interview in diagnosing anxiety disorders. *Arch Gen Psychiatry*. 1988;45:62-63.
5. Hochstim JR. A critical comparison of three strategies of collecting data from households. *J Am Stat Assoc*. 1967;62:976-989.
6. Aneshensel CS, Frerichs RR, Clark VA, Yokopenic PA. Telephone versus in-person surveys of community health status. *Am J Public Health*. 1982;72:1017-1021.
7. Jordan LA, Marcus AC, Reeder LG. Response styles in telephone and household interviewing: a field experiment. *Public Opinion Q*. 1980;44:210-222.
8. Henson R, Roth A, Cannell CF. Personal versus telephone interviews: the effects of telephone reinterviews on reporting of psychiatric symptomatology. In: *National Center for Health Services Research Report Series: Field Experiments in Health Reporting, 1971-1977*. Ann Arbor, Mich: University of Michigan, Survey Research Center, Institute for Social Research; 1977: 205-212.
9. Siemiatycki JA. Comparison of mail, telephone and home interview strategies for household surveys. *Am J Public Health*. 1979;69:238-245.
10. Aneshensel CS, Frerichs RR, Clark VA, Yokopenic PA. Measuring depression in the community, a comparison of telephone and personal interviews. *Public Opinion Q*. 1982;46:110-112.
11. Josephson E. Screening for visual impairment. *Public Health Rep*. 1965;80:47-59.
12. Aneshensel CS, Yokopenic PA. Tests for the comparability of a causal model of depression under two conditions of interviewing. *J Pers Soc Psychol*. 1985;49:1337-1348.
13. Wells KB, Burnam MA, Leake B, Robins LN. Agreement between face-to-face and telephone administered versions of the depression section of the NIMH Diagnostic Interview Schedule. *J Psychiatr Res*. 1988; 22:207-220.
14. Fenig S, Levav I. Demoralization and social supports among Holocaust survivors. *J Nerv Ment Dis*. 1991;179:162-172.
15. Dohrenwend BP, Shrout PE, Egri G, Mendelsohn FS. Nonspecific psychological distress and other dimensions of psychopathology. *Arch Gen Psychiatry*. 1980;37: 1229-1236.
16. Dohrenwend BP, Levav I, Shrout PE. Screening scales from the Psychiatric Epidemiology Research Interview (PERI). In: Weissman MM, Myers JK, Ross CE, eds. *Community Surveys of Psychiatric Disorders*. New Brunswick, NJ: Rutgers University Press; 1986:349-374.
17. Jorm AF, Duncan-Jones P, Scott R. An analysis of the re-test artifact in longitudinal studies of psychiatric symptoms and personality. *Psychol Med*. 1989;19:487-493.
18. Neugebauer R, Kline J, O'Connor P, et al. Depressive symptoms in women in the six months after miscarriage. *Am J Obstet Gynecol*. 1992;166:104-109.